

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF ALABAMA
NORTHWESTERN DIVISION

THOMAS DALE FERGUSON,)
)
Petitioner,)
)
vs.) Case No. 3:09-cv-0138-CLS-JEO
)
)
RICHARD F. ALLEN, Commissioner,)
Alabama Department of Corrections,)
)
)
Respondent.)

MEMORANDUM OPINION

The question addressed in this opinion is whether petitioner, Thomas Dale Ferguson, is categorically excluded from execution by the Supreme Court's opinion in *Atkins v. Virginia*, 536 U.S. 304 (2002), holding that death is not a permissible punishment for an intellectually disabled convict. *Id.* at 321;¹ *see also, e.g., Moore v. Texas*, 137 S. Ct. 1039, 1053 (2017); *Hall v. Florida*, 572 U.S. 701, 713, 722 (2014); *Bobby v. Bies*, 556 U.S. 825, 831 (2009). The issue was presented as a result

¹ When *Atkins* was decided, the condition at issue was referred to as "mental retardation." Within five years, however, the diagnostic terminology employed by mental health professionals was changed to "intellectual disability." *See, e.g.*, Robert A. Schalock *et al.*, *The Renaming of Mental Retardation: Understanding the Change to the Term Intellectual Disability*, 45 Intellectual & Developmental Disabilities 116, 120 (2007) (recognizing that "every individual who is or was eligible for a diagnosis of mental retardation is eligible for a diagnosis of intellectual disability"). Congress also enacted "Rosa's Law" in 2010, which required that all references in federal laws to "mental retardation" be replaced with "intellectual disability." Pub. L. No. 111-256, 124 Stat. 2643, codified as amended at 20 U.S.C. §§ 1140, 1400-01, 7512, and 29 U.S.C. §§ 705, 764, 791 (2010). *See also, e.g., Hall v. Florida*, 572 U.S. 701, 704-05 (2014) (adopting the change in terminology). Thus, the terms "mental retardation" and "mentally retarded" are used in this opinion only when quoting or discussing older authorities.

of granting Ferguson’s Rule 59(e) motion and vacating Part “V.F.” of this court’s prior memorandum opinion, which addressed Ferguson’s claim that he had been improperly denied a hearing on his mental capacity claim.² An evidentiary hearing was held on August 27, 2019. Following consideration of the evidence, pleadings, post-hearing briefs, and additional research, the court enters the following opinion.

I. LEGAL CRITERIA

The Supreme Court’s seminal holding in *Atkins v. Virginia* was based, at least in part, upon the majority’s belief that deficits in the areas of “reasoning, judgment, and control of their impulses” did not allow intellectually disabled criminals to act with the same degree of “moral culpability that characterizes the most serious adult criminal conduct.” 536 U.S. at 306. The Court majority also was not persuaded that the execution of intellectually disabled criminals would advance the deterrent or retributive purposes of the death penalty.³ Consequently, the majority concluded that

² See doc. no. 16 (Memorandum Opinion on Petitioner’s 28 U.S.C. § 2254 Claims for Habeas Corpus Relief), at 146-89; doc. no. 17 (Final Judgment on Petitioner’s 28 U.S.C. § 2254 Claims for Habeas Corpus Relief); doc. no. 18 (Motion to Alter or Amend Judgment); doc. no. 19 (Memorandum Opinion on Petitioner’s Rule 59(e) Motion).

³ The Court elaborated those purposes in, *Hall*, 572 U.S. at 708-09, where the Court observed that:

“[P]unishment is justified under one or more of three principal rationales: rehabilitation, deterrence, and retribution.” *Kennedy v. Louisiana*, 554 U.S. 407, 420 (2008). **Rehabilitation**, it is evident, is not an applicable rationale for the death penalty. See *Gregg v. Georgia*, 428 U.S. 153, 183 (1976) (joint opinion of Stewart, Powell, and Stevens, JJ.). As for **deterrence**, those with intellectual disability are, by reason of their condition, likely unable to make the calculated judgments that are

the execution of such persons was “excessive,” and that the Eighth and Fourteenth Amendments placed “a substantive restriction” upon a state’s power to take the life of an intellectually disabled offender. *Id.* at 321.⁴

Notably, the *Atkins* opinion did not provide definitive procedural or substantive guides for determining when a state prisoner claiming an intellectual disability fell within the protection of the Eighth and Fourteenth Amendments’ prohibition against the imposition of cruel and unusual punishments. Instead, the Court left to the states “the task of developing appropriate ways to enforce the constitutional restriction.” *Atkins*, 536 U.S. at 317 (quoting *Ford v. Wainwright*, 477 U.S. 399, 405, 416-17 (1986) (addressing the execution of insane persons)).

the premise for the deterrence rationale. They have a “diminished ability” to “process information, to learn from experience, to engage in logical reasoning, or to control impulses . . . [which] make[s] it less likely that they can process the information of the possibility of execution as a penalty and, as a result, control their conduct based upon that information.” *Atkins*, 536 U.S., at 320. ***Retributive values*** are also ill-served by executing those with intellectual disability. The diminished capacity of the intellectually disabled lessens moral culpability and hence the retributive value of the punishment. *See id.*, at 319 (“If the culpability of the average murderer is insufficient to justify the most extreme sanction available to the State, the lesser culpability of the mentally retarded offender surely does not merit that form of retribution”). [Alterations in original, boldface emphasis supplied.]

⁴ The Eighth Amendment provides that: “Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted.” U.S. Const., amend. VIII (1791). The prohibition against cruel and unusual punishments was “incorporated” into the Due Process Clause of the Fourteenth Amendment and, thereby, made applicable to the various states by the Supreme Court’s decision in *Robinson v. California*, 370 U.S. 660, 667 (1962). *See also, e.g., Hall*, 572 U.S., at 708 (holding that “No legitimate penological purpose is served by executing a person with intellectual disability. To do so contravenes the Eighth Amendment, for to impose the harshest of punishments on an intellectually disabled person violates his or her inherent dignity as a human being.”) (citing *Atkins*, 536 U.S. at 317).

The Alabama Supreme Court first addressed the criteria for enforcing the constitutional restriction in *Ex parte Perkins*, 851 So. 2d 453 (Ala. 2002), holding that, in order to be classified as mentally retarded (*now described as* “intellectually disabled”), a defendant “must have significantly subaverage intellectual functioning (an IQ of 70 or below), and significant or substantial deficits in adaptive behavior. Additionally, these problems must have manifested themselves during the developmental period (*i.e.*, before the defendant reached age 18).” *Id.* at 456.

The Alabama Supreme Court’s subsequent opinion in *Smith v. State*, 213 So. 3d 239 (Ala. 2007), layered a gloss on the *Perkins* standard, and held that a defendant must exhibit *both* significantly subaverage intellectual functioning *and* significant deficits in adaptive behavior during *three periods* of his or her life: (*i*) before the age of eighteen; (*ii*) on the date of the capital offense; and (*iii*) currently. *Id.* at 248 (citing *Perkins*, 851 So. 2d at 456). *See also Holladay v. Allen*, 555 F.3d 1346, 1353 (11th Cir. 2009) (same).⁵

⁵ The “Defendant with Intellectual Disability Act” was enacted by the Alabama Legislature in August 2009, two years after the decision in *Smith v. State*, and defined the phrase “person with an intellectual disability” as meaning: “A person with [*i*] significant subaverage general intellectual functioning [*ii*] resulting in or associated with concurrent impairments in adaptive behavior and [*iii*] manifested during the developmental period, [*iv*] as measured by appropriate standardized testing instruments.” Ala. Code § 15-24-2(3) (1975) (2018 Replacement Vol.) (alterations supplied). Note that the statutory definition differs from the Alabama Supreme Court’s criteria in two respects: *first*, it *omits* the evidentiary requirement for a defendant’s significantly subaverage intellectual functioning abilities *and* substantial deficits in adaptive behavior to be proven as manifest not just before the age of eighteen (*i.e.*, “during the developmental period”), *but also on the date of the capital offense, as well as currently; and second*, it *adds* the evidentiary requirement for all deficits

II. DISCUSSION

Ferguson bears the burden of proving, by a preponderance of the evidence, that he is intellectually disabled. *See, e.g., Ex parte Carroll*, No. 1170575, 2019 WL 1499322, at *1 (Ala. Apr. 5, 2019). “Intellectual disability is characterized by *significant limitations* both in [a] *intellectual functioning* and in [b] *adaptive behavior* as expressed in conceptual, social, and practical adaptive skills. This disability [c] *originates before age 18.*” American Association on Intellectual & Developmental Disabilities, *Intellectual Disability: Definition, Classification, and Systems of Support* 5 (11th ed. 2010) (emphasis and alterations supplied).⁶

to be “*measured by appropriate standardized testing instruments.*” That diagnostic requirement is the subject of the following section.

⁶ *See also* American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders* (5th ed. 2013) (“DSM-5”), providing that:

Intellectual disability (intellectual developmental disorder) is a disorder with onset during the developmental period that includes both intellectual and adaptive functioning deficits in conceptual, social, and practical domains. The following three criteria must be met:

- A. Deficits in intellectual functions, such as reasoning, problem solving, planning, abstract thinking, judgement, academic learning, and learning from experience, confirmed by both clinical assessment and *individualized standardized intelligence testing.* [Emphasis supplied.]
- B. Deficits in adaptive functioning that result in failure to meet developmental and sociocultural standards for personal independence and social responsibility. Without ongoing support, the adaptive deficits limit functioning in one or more activities of daily life, such as communication, social participation, and independent living, across multiple environments, such as home, school, work, and community.

A. Proof of Significant Limitations in Intellectual Functioning

Proof of the first element of an alleged intellectual disability is best represented by intelligence quotient (“IQ”) scores obtained by the administration of standardized assessment instruments,⁷ such as the “Wechsler Intelligence Scale for Children,” the “Wechsler Adult Intelligence Scale,” or the “Stanford-Binet Intelligence Scales.”

Even though each of the assessment instruments used to measure Ferguson’s IQ at various stages of his life was, on the date of its administration to him, generally considered to be a reliable test, and capable of producing valid scores,⁸ each

C. Onset of intellectual and adaptive deficits during the developmental period.

⁷ The term “assessment instrument” refers to a specific method of acquiring data in the psychological and intellectual assessment of individuals, such as a questionnaire or behavioral observation coding system. *See, e.g.*, Raymond J. Corsini, *The Dictionary of Psychology* 69 (2002) (hereafter, “*Dictionary of Psychology*”). When the adjective “standardized” modifies that term, the combined phrase means that: during the design phase, each instrument was administered to a large, representative sample of the population for which the test was intended to provide reliable, normative data; the validity and reliability of each instrument was established over time by cumulative empirical applications and analyses; and, each individual questionnaire is administered, scored, and interpreted by trained examiners in strict accordance with instructions issued by the test developers. *See, e.g.*, David Wechsler, *WAIS-III Administration and Scoring Manual* (1997); Gale H. Roid, *Stanford-Binet Intelligence Scales, Fifth Edition, Examiner’s Manual* (2003).

⁸ *See, e.g.*, *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 590 n.9 (1993) (“We note that scientists typically distinguish between ‘validity’ (does the principle support what it purports to show?) and ‘reliability’ (does application of the principle produce consistent results?). . . . [O]ur reference here is to *evidentiary reliability* — that is, trustworthiness. . . . In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*.”) (emphasis in original, internal citations omitted).

instrument also contained potential for measurement error.⁹ The Supreme Court's opinion in *Hall v. Florida* recognized that:

The professionals who design, administer, and interpret IQ tests have agreed, for years now, that *IQ test scores should be read not as a single fixed number but as a range*. . . . Each IQ test has a “standard error of measurement,” . . . often referred to by the abbreviation “SEM.” A test’s SEM is a statistical fact, a reflection of the inherent imprecision of the test itself. . . . An individual’s IQ test score on any given exam may fluctuate for a variety of reasons. These include the test-taker’s health; *practice from earlier tests*; the environment or location of the test; the examiner’s demeanor; the subjective judgment involved in scoring certain questions on the exam; and simple lucky guessing. . . .

The SEM reflects the reality that an individual’s intellectual functioning cannot be reduced to a single numerical score. For purposes of most IQ tests, the SEM means that an individual’s score is best understood as a range of scores on either side of the recorded score. The SEM allows clinicians to calculate a range within which one may say an individual’s true IQ score lies. See APA Brief 23 (“SEM is a unit of measurement: 1 SEM equates to a confidence of 68% that the measured score falls within a given score range, while 2 SEM provides a 95% confidence level that the measured score is within a broader range”). A score of 71, for instance, is generally considered to reflect a range between 66 and 76 with 95% confidence and a range of 68.5 and 73.5 with a 68% confidence. See DSM-5, at 37 (“Individuals with intellectual disability have scores of approximately two standard deviations or more below the population mean, including a margin for measurement error (generally +5 points). . . . [T]his involves a score of 65-75 (70 ± 5”). . . . Even when a person has taken multiple tests, each separate score must be assessed using the SEM, and the analysis of multiple IQ scores jointly is a complicated endeavor. . . .

⁹ See, e.g., American Association of Mental Retardation, *Mental Retardation: Definition, Classification, and Systems of Supports* 59 (9th ed. 1992) (hereafter, AAMR, *Mental Retardation*) (“Any trained examiner is aware that all tests contain measurement error; [consequently,] many present scores as confidence bands rather than finite scores.”) (alteration supplied).

Hall, 572 U.S. at 713-14 (citations omitted, emphasis supplied). Thus, the Standard Error of Measurement (“SEM”) accounts for a margin of error of five points, both below and above the test-taker’s IQ score on the standardized assessment instrument at issue. *Ledford v. Warden, Georgia Diagnostic & Classification Prison*, 818 F.3d 600, 640 (11th Cir. 2016).¹⁰

1. IQ scores found in Ferguson’s public school records

During November of 1979, when Ferguson was six years of age, he obtained a Full-Scale IQ score of 77 on the Stanford-Binet Intelligence Scale.¹¹

Six years later, when Ferguson was on the threshold of the 1985-86 school year, he was evaluated at the request of his mother for special educational services because of a noted “lack of academic progress, suspected learning disability, deficient reading skills, and deficient handwriting skills.” Doc. no. 39-5 (Petitioner’s Exhibit 5), at 21 (Morgan County School System Confidential Student Evaluation). The “Wechsler Intelligence Scale for Children-Revised (WISC-R)” assessment instrument was administered on August 6, 1985, and he achieved a Verbal IQ score

¹⁰ See also, e.g., *Walker v. True*, 399 F.3d 315, 322 (4th Cir. 2005) (“IQ tests have a measurement error of plus or minus five points.”).

¹¹ See doc. no. 39-5 (Petitioner’s Exhibit 5), at 35, 37. **Note well:** *The page numbers of all documents filed in this court will be those imprinted at the top of each page by the CM/ECF automated filing system, and not necessarily those of counsel.* This can be confusing when citing such pleadings as doc. no. 47 (Petitioner’s Brief), which begins with an unnumbered cover page, followed by a “Table of Contents” page numbered with a lower-case Roman numeral. As a result, the first page bearing Arabic numeral “1” was numbered “Page 3” by the CM/ECF filing system.

of 74, a Performance IQ score of 71, and a Full-Scale IQ score of 71. *Id.* at 22.¹² Significantly, however, even though Ferguson “did not appear to be challenged by the more difficult items on the test,” the test administrator noted that he “gave up easily on both verbal and non-verbal items.”¹³ After reviewing the test scores, the school system concluded that Ferguson was eligible for special services as an “educationally mentally handicapped” student.¹⁴ Such children typically were segregated from the main student body and placed in “a self-contained classroom with about six or seven other students.”¹⁵

During March of 1988, when Ferguson was fifteen and in the second semester of his 8th Grade, 1987-88 school year, school system policy required that he be re-

¹² See also doc. no. 39-5 (Petitioner’s Exhibit 5), at 37 (August 12, 1985 Morgan County Board of Education “Eligibility Justification” stating that Ferguson “continues to be eligible for special services as Educationally Mentally Handicapped. . . . Dale’s re-evaluation indicates that his overall functioning is within the EMH range. A previous Binet (11/15/79) yielded an IQ score of 77 which is considered to be consistent with his current WISC-R score of 71.” (ellipsis supplied)).

¹³ *Id.* at 22. This court rejects Ferguson’s contention that the comments of the examiner that are quoted in text cannot be used to conclude that his low score on this test was due to malingering, because the examiner also stated that Ferguson was “cooperative during the entire session,” and that his scores were “considered to be valid.” Doc. no. 47 (Petitioner’s Brief), at 17 (citing doc. no. 39-5 (Petitioner’s Exhibit 5), at 22).

¹⁴ Doc. no. 39-5 (Petitioner’s Exhibit 5), at 36-37 (August 12, 1985 justification for providing Ferguson special services as an educationally mentally handicapped student). Terms such as “Educationally Mentally Handicapped,” “Educable Mentally Impaired,” “Trainable Mentally Handicapped,” and “Educable Mentally Retarded” were virtually sonorous descriptions of similar conditions. *See, e.g., Dictionary of Psychology* 312.

¹⁵ See doc. no. 46 (Hearing Transcript), at 127 (where the State’s expert, Dr. Glen King, testified that: “Typically in Alabama if you’re in an [educable mentally retarded] class, you’re in a self-contained classroom with about six or seven other students and you stay there all day”).

evaluated for continued receipt of the special services provided to educationally mentally handicapped students.¹⁶ The WISC-R was re-administered on March 1, 1988, and Ferguson achieved a Verbal IQ score of 87, a Performance IQ score of 88, and a Full-Scale IQ score of 87 — a decided, sixteen-point improvement over his performance on the same assessment instrument nearly three years before.¹⁷ (During the hearing held in this court, Ferguson’s own expert, Dr. Robert D. Shaffer, confirmed that Ferguson’s scores on six of the ten subtests comprising the WISC-R improved on the 1988 re-test, and that he did not score two standard deviations below the mean on any subtest.¹⁸) The test administrator recorded that Ferguson’s full-scale score fell “within the low average range of intellectual functioning.”¹⁹ Consequently, he was moved into classes for students classified as “learning disabled.”²⁰ Such students were not segregated from the larger student body, but received additional assistance with certain subjects.²¹

¹⁶ See doc. no. 39-5 (Petitioner’s Exhibit 5), at 51 (“Dale [Ferguson] is currently enrolled in the special education program *and is due the required three year re-evaluation for continued placement in that program.*”) (alteration and emphasis supplied).

¹⁷ *Id.* at 52.

¹⁸ Doc. no. 46 (Hearing Transcript), at 56-57.

¹⁹ See doc. no. 39-5 (Petitioner’s Exhibit 5), at 52.

²⁰ *Id.* at 60.

²¹ See doc. no. 46 (Hearing Transcript), at 127-28 (where the State’s expert, Dr. Glen King, testified that students classified as “learning disabled” are “mainstreamed,” meaning that they are “in regular classes all day,” but receive “special resources” (e.g., “additional help with reading, writing”), and typically are provided an “Individualized Education Program” which allows “extra time to take tests, or you have [test] items . . . read to you, or things of that nature”) (alteration

2. State court IQ assessments

While awaiting trial in state court, Ferguson was referred by the court to Dr. C. Van Rosen for a psychological evaluation to assist in determining his competency to stand trial, as well as his mental state at the time of the offense.²² Dr. Rosen administered a “Wechsler Adult Intelligence Scale–Revised (WAIS-R)” assessment instrument during December of 1997, and recorded that Ferguson achieved a Verbal IQ score of 76, a Performance IQ score of 66, and a Full-Scale IQ score of 69, which Dr. Rosen characterized as

technically plac[ing] him in the very high range of the mildly retarded level. *However, it was considered quite apparent that he did not attempt to make a good effort in this test, giving up readily on many items and seemingly not trying as hard as possible. The defendant's intellectual functioning is consequently considered to be within the higher portion of the borderline range of abilities.*

Doc. no. 41-9 (Respondent’s Exhibit 5), at 7 (Rosen Report) (alteration and emphasis supplied). Dr. Rosen subsequently testified during the state-court trial that: he had seen no signs of mental retardation during his evaluation of Ferguson; he was certain that Ferguson did not put forth his full effort during administration of the WAIS-R;

supplied).

²² See doc. no. 41-9 (Respondent’s Exhibit 5), at 1 (Rosen Report).

and that, if Ferguson “had really tried[,] he would have scored probably in the middle 70’s . . . perhaps a little higher.”²³

Ferguson was referred by his defense attorneys to Dr. James F. Chudy for an independent psychological evaluation prior to trial. Dr. Chudy performed his examination in June of 1998. Although he did not include a specific IQ score in his written report of evaluation,²⁴ he concluded, after administering a battery of standardized personality tests, that none of the tests revealed “organic problems,”²⁵ but all in combination demonstrated that Ferguson was “functioning in the borderline range” of intelligence, in “the area between low average intelligence and mental retardation.”²⁶

3. Recent IQ assessments

Ferguson’s most recent IQ scores were derived from testing performed in preparation for the evidentiary hearing in this court.

a. Ferguson’s psychologist — Dr. Robert D. Shaffer

²³ Doc. no. 41-9 (Respondent’s Exhibit 6), at 25 (Rosen Trial Testimony).

²⁴ See doc. no. 41-9 (Respondent’s Exhibit 7), at 43-51 (Chudy Written Report).

²⁵ See doc. no. 41-9 (Respondent’s Exhibit 8), at 57-58 (Chudy Trial Testimony: “I found some memory problems. They were fairly consistent with his low intelligence, but no significant organic problems.”).

²⁶ *Id.* at 58 (Chudy Trial Testimony).

Robert Daniel Shaffer, who holds a Ph. D. in clinical psychology and has specialized training in neuropsychology and forensic psychology, was retained by Ferguson's attorneys to evaluate his "cognitive and intellectual functions, and his adaptive behavior in light of current court rulings pertaining to Intellectual Disability and death penalty eligibility."²⁷ He met with Ferguson for a total of 13.8 hours over three days.²⁸ He also interviewed Ferguson's mother, Mrs. Betty Ferguson, for an additional 4.2 hours.²⁹ He concluded that Ferguson demonstrated "significantly sub-average intellectual ability,"³⁰ and recorded that his condition was

most evident in the profile of neuropsychological functions. The DSM-V states that this profile of abilities is to be more strongly weighed than single IQ numerical scores, owing to the imprecision of IQ tests themselves. This profile of neurocognitive abilities reveals significant sub-average deficits in intellectual ability. IQ scores obtained in this evaluation are consistent with this finding.

Doc. no. 39-1 (Petitioner's Exhibit 1), at 16.

Dr. Shaffer's opinion relied, at least in part, upon the Full-Scale IQ score of 77 achieved by Ferguson during the September 18, 2017 administration of a "Wechsler

²⁷ Doc. no. 39-1 (Petitioner's Exhibit 1), at 1.

²⁸ Dr. Shaffer's evaluations of Ferguson occurred on September 18 and 19, 2017, and March 22, 2018. *Id.* Notably, Shaffer's last evaluation was conducted the month after Ferguson was examined by the State's psychologist.

²⁹ Dr. Shaffer interviewed Mrs. Betty Ferguson on January 29, 2018 (*id.*) in an attempt to assess Ferguson's adaptive functions "from her reference point" during the periods she and Ferguson lived together. *See* doc. no. 46 (Hearing Transcript), at 12.

³⁰ Doc. no. 39-1 (Petitioner's Exhibit 1), at 9.

Adult Intelligence Scale–Fourth Edition (WAIS-IV)” assessment instrument. Adjustment of that score for the “Flynn effect”³¹ reduced it to 73.4.³² After applying the standard error of measurement to the Flynn-effect-adjusted score, Dr. Shaffer testified that there was a 95% probability that Ferguson’s “true IQ” lay in a range between 69.4 and 78.4, and a 99% probability that his true IQ lay somewhere in a range between 67.9 and 78.9.³³ Thus, according to Dr. Shaffer, Ferguson’s intellectual functioning is significantly substandard, because his true IQ *could be* less than 70.

Dr. Shaffer also administered several neuropsychological tests in an attempt to assess Ferguson’s executive, verbal, visual-processing, organizational, memory, and fine-motor functions. Not all of the tests demonstrated substantial impairments, but Dr. Shaffer concluded that the totality of all test results were consistent with his

³¹ The “Flynn effect” is a term that refers to the finding that average IQ scores have increased steadily in the United States and European nations for many decades. The concept is discussed more thoroughly in Part II.A.4.a of this opinion, *infra*.

³² See doc. no. 46 (Hearing Transcript), at 26-27. Dr. Shaffer said that, when adjusting Ferguson’s Full-Scale IQ score for the “Flynn effect,” he began by calculating the number of years that had elapsed between the date on which the WAIS-IV assessment instrument had been standardized, or “normed,” and the date on which he administered the test to Ferguson (September 18, 2017). He testified that difference was 10.2 years. *Id.* at 27. (10.2 x 0.3 = -3.6 deducted from 77 = IQ score of 73.4 following adjustment for the Flynn effect). However, as discussed in the textual paragraph accompanying notes 64 and 65, *infra*, this court concludes that Dr. Shaffer’s adjustment calculations require revision because he miscalculated the number of years that elapsed between the date on which the WAIS-IV assessment instrument was published and the date on which he administered it to Ferguson.

³³ See *id.* at 27-28.

opinion that Ferguson has significant limitations in his ability to function intellectually.³⁴

b. The State's psychologist — Dr. Glen D. King

Glen David King, who holds a J.D. degree and Ph. D. in clinical and forensic psychology, was retained by the Alabama Attorney General's office “to primarily determine the intellectual ability of Mr. Ferguson.”³⁵ Dr. King met with him for five-and-a-half hours on February 13, 2018 — five months after Dr. Shaffer had completed the second of his three meetings with Ferguson — and administered six tests, “just to make sure” that he employed at least one assessment instrument that had not previously been administered by Dr. Shaffer.³⁶ Even so, Dr. King’s opinions were principally grounded upon the results obtained from his re-administration of the

³⁴ See *id.* at 29-32. See also doc. no. 39-1 (Petitioner’s Exhibit 1), at 5-6 & 7-13 (where Shaffer summarizes the neuropsychological tests administered to Ferguson and the results obtained from each).

³⁵ Doc. no. 41-1 (Respondent’s Exhibit 1), at 1.

³⁶ *Id.* at 3; Doc. no. 46 (Hearing Transcript), at 122. The assessment instruments administered by Dr. King were: the “Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV)”; the “Stanford Binet Intelligence Scale–Fifth Edition (SB-5)”; the “Wide Range Achievement Test–Fifth Edition (WRAT-5)”; the “Adaptive Behavior Assessment System–Third Edition (ABAS-3)”; the “Independent Living Scales (ILS)”; and, a test he referred to as the “Shipley-2.” Doc. no. 41-1 (Respondent’s Exhibit 1), at 2, 5-7. (“Shipley-2” is an abbreviation of the “Shipley-Institute for Living Scale for Measuring Intellectual Impairment–Second Edition (Shipley-2).” See <https://www.stoeltingco.com/shipley-institute-of-living-scale-second-edition-shipley-2.html> (last visited Apr. 6, 2020). The test is “used for determining pathological deterioration of intelligence. A conceptual quotient is obtained and scores below 100 are purported to indicate mental impairment due to age, disease or injury.” *Dictionary of Psychology* 899.

“Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV)” test,³⁷ as well as his unique administration of a “Stanford Binet Intelligence Scale–Fifth Edition (SB-5)” assessment instrument.³⁸ Ferguson obtained Full-Scale IQ scores of 85 on the WAIS-IV, and 84 on the SB-5.³⁹

Dr. King testified that Ferguson’s Full-Scale IQ score of 85 on the WAIS-IV placed him on the boundary between two categories of intellectual ability: that is, his IQ score could be classified as lying *either at* the high end of the “borderline range” of intellectual ability (*i.e.*, IQ scores between 70 and 84), *or at* the low end of the “average range” of intellectual ability (*i.e.*, IQ scores from 85 to 115).⁴⁰ After adjusting Ferguson’s 85 score for the standard error of measurement, Dr. King stated that an evaluator could be 99% confident that Ferguson’s Full-Scale IQ score on the

³⁷ Dr. King described the WAIS-IV as a test that had “ten subtests associated with it. Those ten subtests are divided into four different domains: verbal comprehension index, perceptual reasoning index, working memory index, and processing speed index. We get a score on each one of those indices, and then they’re all combined to give us a full scale IQ score.” Doc. no. 46 (Hearing Transcript), at 117-18. Scores on the WAIS-IV subtests, as well as the full-scale score, average 100 (*the mean*), with a standard deviation of 15, which means that a Full-Scale IQ score of 85 would be “one standard deviation below the mean.” *Id.* at 118.

³⁸ Dr. King described the “Stanford Binet Intelligence Scale–Fifth Edition (SB-5)” as being “an individually administered intelligence test. It has no items that overlap with the Wechsler Adult Intelligence Scale, but [it] is also considered to be one of the two gold standard IQ tests for measuring IQ and school placement. It is – it has also an average of a hundred and a standard deviation of 15.” *Id.* at 120 (alteration supplied).

³⁹ See doc. no. 41-1 (Respondent’s Exhibit 1), at 5-6 & 8.

⁴⁰ See doc. no. 46 (Hearing Transcript), at 119, lines 12-14 (testifying that IQ scores between 70 to 84 are classified in “the borderline range,” while IQ scores from 85 to 115 are classified as “average” intelligence).

WAIS-IV would lie somewhere between 81 and 89,⁴¹ a span that straddled both the “borderline” and “average” ranges of intellectual functioning.

Ferguson’s Full-Scale IQ score of 84 on the “Stanford Binet Intelligence Scale–Fifth Edition (SB-5)” assessment instrument placed him “at the high end of the borderline range” of intellectual functioning.⁴² The closeness of Ferguson’s scores on the WAIS-IV and SB-5 — Dr. King characterized them as being in “absolute agreement with each other”⁴³ — indicated that the tests had accurately measured Ferguson’s intellectual abilities. Dr. King added that, while a test subject can “always malingering, or not put forth good effort and have a reduced score” on an IQ test, it is not possible to “fake good,” and “the highest score . . . would be an indication of your best performance.”⁴⁴

Dr. King also reviewed the IQ scores found in Ferguson’s public school records,⁴⁵ and testified that the full-scale score of 87 attained on the “Wechsler Intelligence Scale for Children–Revised (WISC-R)” in March of 1988, when he was fifteen, was “in good agreement with what I found on all of the tests I gave him.”⁴⁶

⁴¹ *Id.* at 119-20.

⁴² *Id.* at 120-21.

⁴³ *Id.* at 122.

⁴⁴ *Id.* at 123.

⁴⁵ Those scores were summarized in Section II.A.1 of this opinion, *supra*.

⁴⁶ Doc. no. 46 (Hearing Transcript), at 125. King did not believe that the low, Full-Scale IQ score of 71 obtained by Ferguson on the WISC-R administered in July 1985, when he was twelve

Based upon such considerations, Dr. King testified that Ferguson’s intellectual functioning abilities were not “significantly subaverage.” Instead, in Dr. King’s opinion, Ferguson’s IQ scores placed him either at the high end of the borderline range, or at the low end of the average range, of intellectual functioning.⁴⁷

4. The parties’ contentions:

a. Adjustment for the Flynn effect

Ferguson contends that, “contrary to established procedures in the authoritative texts, Dr. King failed to quantitatively or qualitatively account for the shifting norms” of his IQ tests by adjusting his scores for the “Flynn effect”⁴⁸ — a term that refers to the discovery in 1984 by political scientist James R. Flynn that average IQ scores had

years old (and indicating that he was in the borderline range of intellectual functioning), cast doubt upon the accuracy of the IQ scores obtained in February 2018, because the notation of the 1985 examiner indicated that Ferguson “gave up easily, [and] was not necessarily putting forth his best effort”: in other words, he could have done better. *Id.* at 124 (alteration supplied). King added that Ferguson’s school records indicated that he suffered from a “learning disability,” as opposed to an intellectual disability. *See id.* at 125-26.

⁴⁷ *Id.* at 129 (“Based on all the tests together, he’s somewhere right at the high borderline to below average.”). King also noted that, with but one exception — that being the WISC-R test administered in July 1985 that is addressed in the preceding footnote — Ferguson scored somewhere between 70 and 85 or 86 on every IQ test administered over the course of his life. *Id.* at 130.

Dr. King also reviewed Ferguson’s prison records, spanning an eighteen year period from 1998 to 2016, and found that he “was in segregation on multiple occasions and had to be reviewed. *In every one of those occasions the rating was again normal intelligence* [even though] there were boxes to check off for borderline or disabled. So there were no indications [of intellectual disabilities] from that.” *Id.* at 128 (emphasis and alterations supplied). Finally, King found a letter written by Ferguson in his prison files, requesting a transfer to another penal facility closer to family, and King described it as notable in several respects: *i.e.*, it was divided into “paragraphs,” “neatly written,” “punctuated and spelled correctly,” and demonstrated “good handwriting.” *Id.* at 129.

⁴⁸ Doc. no 47 (Petitioner’s Brief), at 16.

increased steadily in the United States and Europe for more than forty years at an average rate of 3.0 points a decade, or 0.3 points annually.⁴⁹

Dr. King based his failure to make such an adjustment on his opinion that the Flynn effect is only “a theory” that has “not been completely validated.”⁵⁰ Moreover, Dr King stated that, when standardized assessment instruments are administered to the same subject over a long period of time, and the IQ scores generally

agree with each other, it actually increases your expectation that the score that you’re getting is accurate. So the idea is that if you keep getting scores in the 70s and 80s, the scores are in the 70s and 80s. You don’t need to alter them from that point.

Doc. no. 46 (Hearing Transcript), at 137-38.

The disagreement between Drs. Shaffer and King reflects the Eleventh Circuit’s observation that there is “no consensus about the Flynn effect among experts or among the courts.” *Raulerson v. Warden*, 928 F.3d 987, 1008 (11th Cir. 2019); *see also Thomas v. Allen*, 607 F.3d 749, 757-58 (11th Cir. 2010) (same); *Ledford*, 818

⁴⁹ James R. Flynn, *The Mean IQ of Americans: Massive Gains 1932 to 1978*, 95 *Psychol. Bull.* 29 (1984); *see also* James R. Flynn, *Massive IQ Gains in 14 Nations: What IQ Tests Really Measure*, 101 *Psychol. Bull.* 171 (1987). Flynn’s studies revealed a 13.8-point increase in IQ scores in the standardization samples of successive versions of Stanford-Binet and Wechsler intelligence tests between 1932 and 1978, amounting to a 0.3-point increase per year, or approximately 3 points per decade. More recently, the Flynn effect was supported by calculations of IQ score gains between 1972 and 2006 for different normative versions of the Stanford-Binet (SB), Wechsler Adult Intelligence Scale (WAIS), and Wechsler Intelligence Scale for Children (WISC). The average increase in IQ scores per year was 0.31, which was consistent with Flynn’s earlier findings. James R. Flynn, *Tethering the Elephant: Capital Cases, IQ, and the Flynn Effect*, 12 *Psycho. Pub. Pol’y & L.* 170, 177-78 (2006).

⁵⁰ *See* doc. no. 46 (Hearing Transcript), at 130-31.

F.3d at 635-37 (explaining the divergent approaches to the Flynn effect taken by other circuits).

Any attempt at reconciliation of that disagreement must begin with recognition of the fact that intelligence assessment instruments are rarely revised, sometimes no more frequently than every fifteen or twenty years. For example, the Wechsler Adult Intelligence Scale was developed in 1955,⁵¹ but only updated in 1981,⁵² 1997,⁵³ and, most recently, 2008.⁵⁴ Each version had to be “standardized,” meaning that during the design phase, each test instrument was administered to a large, representative sample of the population for whom the test was intended to provide reliable, normative data. To promote accuracy and avoid bias, the standardization sample of individuals had to be representative of all potential test subjects, which meant that test developers first had to identify and track key census data variables, such as age, gender, race, socio-economic status, educational attainments, and geographic

⁵¹ The WAIS was released by David Wechsler in February 1955, as a revision of the “Wechsler–Bellevue Intelligence Scale (WBIS)”: a battery of tests published by Wechsler in 1939. See *Wechsler Adult Intelligence Scale*, Wikipedia, https://en.wikipedia.org/wiki/Wechsler_Adult_Intelligence_Scale (last visited Apr. 20, 2020).

⁵² The “Wechsler Adult Intelligence Scale–Revised (WAIS-R),” a revised form of the WAIS, was released in 1981. It consisted of six verbal and five performance subtests, from which Verbal IQ, Performance IQ, and Full-Scale IQ scores were obtained. *Id.*

⁵³ The “Wechsler Adult Intelligence Scale–III (WAIS-III),” a subsequent revision of the WAIS and the WAIS-R, was released in 1997. It provided scores for Verbal IQ, Performance IQ, and Full-Scale IQ, along with four secondary indices (Verbal Comprehension, Working Memory, Perceptual Organization, and Processing Speed). *Id.*

⁵⁴ The current version of the test, the “Wechsler Adult Intelligence Scale–IV (WAIS-IV),” was released in 2008. *Id.*

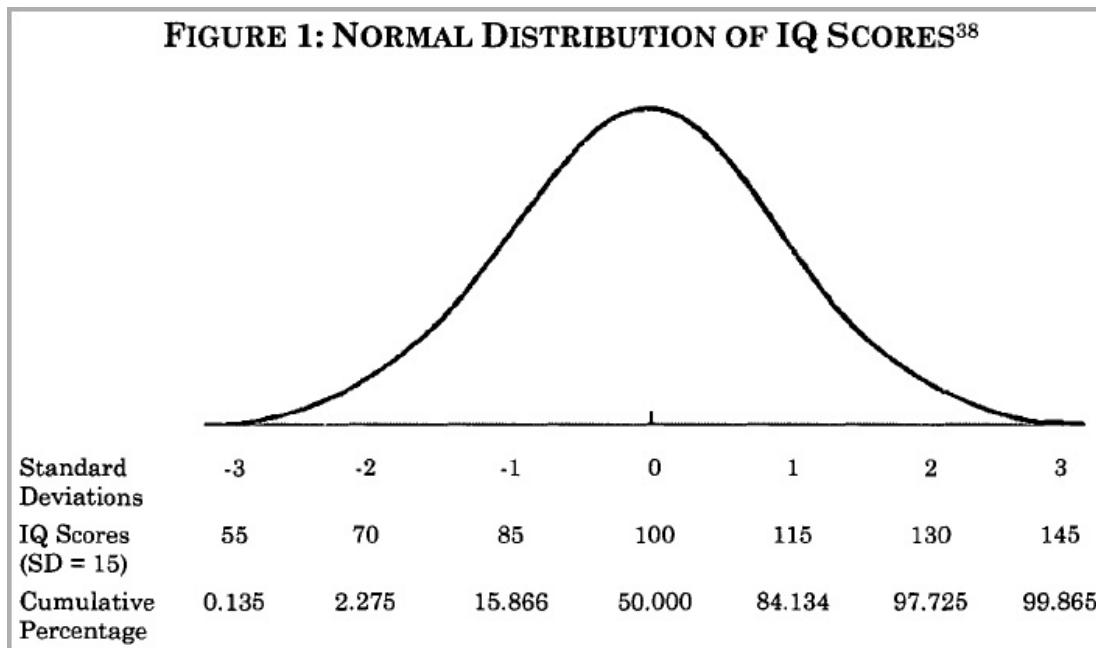
residence. Test developers then selected a “standardization,” or “normative,” sample of individual test subjects matching the census proportions. For example, the current, fourth edition of the Wechsler Adult Intelligence Scale was standardized on a normative sample of 2,200 U.S. citizens ranging in age from 16 to 90, and stratified by gender, education level, ethnicity, and geographic region of residence.⁵⁵ In like manner, the current “Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V),” was released in 2014, and had been standardized on a normative sample of 2,200 children between the ages of 6 and 16 years & 11 months.⁵⁶ Following administration of the test to each member of the standardization sample, a graph was created from the scores achieved by the test subjects. The graph is in the form of a normal distribution, and the mean or median raw score of the normative sample:⁵⁷

⁵⁵ *Id.*

⁵⁶ See *Wechsler Intelligence Scale for Children*, Wikipedia, https://en.wikipedia.org/wiki/Wechsler_Intelligence_Scale_for_Children (last visited Apr. 20, 2020).

⁵⁷ The terms “average,” “mean,” and “median” can be confusing. The **average** and **arithmetic mean** are calculated by adding together all numerical values in a data set, and dividing that sum by the number of terms in the set. For example, if there were nine numerical values in a data set (e.g., 30, 56, 65, 70, 84, 90, 95, 110, 130), the sum of all added together would be 730, and 730 divided by 9 yields 81.11, which is the “average,” or “arithmetic mean.” In contrast, the **median** is the middle numerical value in an ordered set of data. If the set has *an odd number of data points*, then regardless of whether the numerical values are ranked from lowest to highest (e.g., 30, 56, 65, 70, 84, 90, 95, 110, 130), or from highest to lowest (e.g., 130, 110, 95, 90, 84, 70, 65, 56, 30), the “median” is the numerical value in the middle: here, 84. On the other hand, if the ordered set contains *an even number of data points*, then the “median” is the “arithmetic mean” (or average) of the two data points in the middle: e.g., in an ordered data set of ten numbers ranked from lowest to highest (30, 56, 65, 70, 75, 84, 90, 95, 110, 130), the sum of the middle numbers 75 + 84 is 159, which when divided by 2 yields the “median” of 79.5. The same would be true if the ordered set were ranked from highest to lowest: e.g., 130, 110, 95, 90, 84, 75, 70, 65, 56, 30.

always is arbitrarily defined as the number “100”; always is located at the apex of the symmetric, “bell-shaped” curve; and always means that a score of 100 represents average performance on the IQ test. The standard deviation is defined as fifteen points above and below the median. Figure 1 displays a normal distribution of IQ scores, illustrating how scores are distributed based on standard deviations:⁵⁸



On any of the Wechsler IQ assessment instruments, with a mean or median of 100, and a fifteen-point standard deviation, one standard deviation from the mean, encompassing about two-thirds of all test takers, results in a range of scores between 85 and 115 (fifteen points below and above the mean score of 100). Two standard

⁵⁸ The following Figure was copied from an excellent student note authored by Geraldine W. Young, and published in the March 2012 edition of the Vanderbilt University Law Review: *i.e.*, Geraldine W. Young, *A More Intelligent and Just Atkins: Adjusting for the Flynn Effect in Capital Determinations of Mental Retardation or Intellectual Disability*, 65 Vand. L. Rev. 615, 622 (2012).

deviations from the mean, encompassing about ninety-five percent of all test takers, results in a range of scores between 70 and 130 (thirty points below and above the mean score). A score of 130 on Wechsler tests often operates as the lower threshold for a classification of “giftedness,” while a score of 70 to 75 often marks the upper threshold for “intellectual disability” (or, as it formerly was called, “mental retardation”).⁵⁹

The foregoing description of the process for standardizing an IQ assessment instrument underscores that it is a lengthy and expensive process. For such reasons, test publishers do not frequently re-standardize tests. Moreover, given the effort and expense required to replace aging assessment instruments, administrators often continue to use tests standardized many years before — a fact that results in the administration of IQ tests a decade or more after the publishers originally normed the standardization sample.⁶⁰

⁵⁹ See, e.g., *Atkins*, 536 U.S. at 309 n.5 (“It is estimated that between 1 and 3 percent of the population has an IQ between 70 and 75 or lower, which is typically considered the cutoff IQ score for the intellectual function prong of the mental retardation definition.”) (citing 2 Kaplan & Sadock’s Comprehensive Textbook of Psychiatry 2952 (B. Sadock & V. Sadock eds., 7th ed.2000)).

⁶⁰ The foregoing discussion of the manner of creating IQ assessment instruments is based upon Young, *A More Intelligent and Just Atkins*, 65 Vand. L. Rev. at 621-23. See also Alan S. Kaufman, *IQ Testing 101* 107, 125, 130 (2009); William E. Benet, *Genius: An Overview*, Assessment Psychol. Online (Jan. 2005), <http://www.assessmentpsychology.com/genuis2.htm>; *IQ Classifications*, Assessment Psychol. Online, <https://www.assessmentpsychology.com/iqclassifications.htm> (last visited May 1, 2020).

Dr. Flynn's original, 1984 study documented IQ gains among American citizens over a period of more than forty years, from 1932 through 1978.⁶¹ He used data collected from seventy-three previous studies, with a combined total of almost 7,500 subjects to whom various Stanford-Binet and Wechsler IQ assessment instruments had been administered. He discovered that, on average, the subjects scored higher on earlier-normed IQ tests than on later-normed IQ tests.⁶²

Flynn then calculated IQ gains by (1) measuring the difference between the subjects' mean scores on the two tests — the earlier and later tests, normed at different times but taken around the same time — and (2) dividing that difference by the number of years that had passed between the norming of the earlier test and the norming of the later test, resulting in a figure around 0.3 points per year.

From his findings, Flynn concluded that the IQ gains reflected the obsolete, outdated norms of the earlier IQ tests, as compared to later, more recently normed tests. *Flynn defined obsolete norms as “simply ones that are earlier and easier than later norms.”* Given the relative nature of IQ scores and the observed IQ gains over time, *when a person takes an IQ test in 2010 and that test’s norms are based on a standardization sample from 1980, the person receives an inflated score because the score is based on the weaker performance of the 1980 sample, rather than the performance of the person’s peers in 2010.* In other words, a person's IQ score on an earlier test with obsolete norms may be above average, while the same performance may be average or below average on a later test with updated norms. Psychologist and IQ-test developer Alan S. Kaufman explains this phenomenon with an analogy: while a runner's time may have won a track meet twenty years ago, the same time may not even qualify for a meet today.

⁶¹ James R. Flynn, *The Mean IQ of Americans: Massive Gains 1932 to 1978*, 95 Psychol. Bull. 29 (1984).

⁶² *Id.* at 32.

Geraldine W. Young, *A More Intelligent and Just Atkins: Adjusting for the Flynn Effect in Capital Determinations of Mental Retardation or Intellectual Disability*, 65 Vand. L. Rev. 615, 624-25 (2012) (footnotes omitted, emphasis supplied).

For such reasons, this court adjusted the IQ scores of a habeas petitioner in accordance with the Flynn effect in a previous *Atkins* appeal. *Thomas v. Allen*, 614 F. Supp. 2d 1257, 1275-81 (N.D. Ala. 2009), *affirmed* 607 F.3d 749, 753 (11th Cir. 2010) (holding that an evaluator “may also consider the ‘Flynn effect,’ a method that recognizes the fact that IQ test scores have been increasing over time”). Nothing that has been presented in this case by either Dr. King or the State’s attorneys persuades this court that it should not do so again.⁶³

⁶³ The *User’s Guide* to the tenth edition of the American Association on Intellectual and Development Disabilities’ treatise on *Mental Retardation Definition, Classification and Systems of Supports* recommends that clinicians take the Flynn effect into account when performing evaluations in less than optimal circumstances (e.g., the legal and physical constraints of a maximum-security prison environment). Specifically, the *Guide* directs diagnosticians to:

Recognize the “Flynn Effect.” In his study of IQ tests across populations, Flynn (1984, 1987, 1989) discovered that IQ scores have been increasing from one generation to the next in all 14 nations for which IQ data existed. This increase in IQ scores over time has been dubbed the Flynn Effect. Flynn reported a greater increase in the Wechsler Performance IQ, which is more heavily loaded on fluid abilities than on the Wechsler Verbal IQs. On average, the Full-Scale IQ increases by approximately 0.33 points for every year elapsed since the test was normed (Flynn, 1999). The main recommendation resulting from this work is that all intellectual assessments must use a reliable and appropriate individually administered intelligence test. In cases of tests with multiple versions, the most recent version with the most current norms should be used at all times. In cases where a test with aging norms is used, a correction for the age of the norms is warranted. For example, if the Wechsler Adult Intelligence Scale (WAIS-III, 1997) was used to assess an individual’s IQ in July 2005, the population mean on the WAIS-III was set at 100 when it was originally normed in 1995 (published in 1997). However, based on

Having said that, however, does not mean that this court unequivocally endorses Dr. Shaffer's conclusions. His adjustment calculations require revision because he miscalculated the number of years that elapsed between the date on which the WAIS-IV assessment instrument was published and the date on which he administered it to Ferguson. He computed that difference as 10.2 years, which he then multiplied by the factor 0.3, the rate at which Flynn predicted outdated IQ norms increase each year. However, the WAIS-IV was released in August 2008,⁶⁴ and the difference between that date and the September 18, 2017 date on which Dr. Shaffer administered the test to Ferguson was nine years and one-and-a-half months, or 9.08

Flynn's data, the population mean on the Full-Scale IQ raises roughly 0.33 points per year; thus, the population mean on the WAIS-III Full-Scale IQ corrected for the Flynn Effect would be 103 in 2005 (9 years x 0.33 = 2.9). Hence, using the "at least two standard deviations below the mean" (Luckasson et al., 2002), the approximate Full-Scale IQ cutoff would be approximately 73 (plus or minus the standard error of measurement). Thus the clinician needs to use the most current version of an individually administered test of intelligence and take into consideration the Flynn Effect as well as the standard error of measurement when estimating an individual's true IQ score.

American Association on Intellectual and Developmental Disabilities, *User's Guide: Mental Retardation Definition, Classification and Systems of Supports* 20-21 (10th ed. 2006) (hereafter, AAIDD, *User's Guide*). See also, e.g., American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders* 37 (5th ed. 2013) (hereafter "DSM-5") ("Factors that *may* affect [intelligence] test scores include practice effects and the 'Flynn effect' (i.e., overly high scores due to out-of-date test norms)") (emphasis and alteration supplied); Marc J. Tassé et al., *The Relation Between Intellectual Functioning and Adaptive Behavior in the Diagnosis of Intellectual Disability*, 54 Intell. & Developmental Disabilities 381, 382 (2016) ("All other sources of measurement error, such as the *Flynn effect* and *practice effects*, should also be considered when interpreting test results.") (emphasis supplied).

⁶⁴ David Wechsler, *Wechsler Adult Intelligence Scale: Technical and Interpretative Manual* 22 (4th ed. 2008).

years. Multiplying 9.08 by the Flynn factor of 0.3 produces a product of 2.7 which, when subtracted from Ferguson’s Full-Scale IQ score of 77, yields an adjusted IQ score of 74.3 — not, as Dr. Shaffer testified, 73.4.⁶⁵

b. Adjustment for the standard error of measurement

Ferguson also accuses Dr. King of ignoring proper diagnostic practices by reporting his IQ scores as a single fixed number, as opposed to a range of numbers accounting for the statistical error of measurement.⁶⁶ That criticism is not entirely fair, however, in view of King’s testimony that, after adjusting Ferguson’s Full-Scale IQ score of 85 on the WAIS-IV for the standard error of measurement, an evaluator could be 99% confident that his “true IQ” would lie somewhere between 81 and 89:⁶⁷ a range that straddled both the “borderline” and “average” ranges of intellectual functioning. The more pertinent criticism is addressed in the following section.

c. Adjustment for the “practice effect”

⁶⁵ The only calculations that come close to Dr. Shaffer’s testimony must begin with March 2007: the date during which data began to be collected from individuals who ultimately comprised the normative sample for the WAIS-IV that was published in August 2008. The difference between March 2007 and Shaffer’s September 18, 2017 administration was ten years and seven months, or 10.583 years, which — when multiplied by the Flynn factor of 0.3 — produces a product of 3.1749. Subtracting that product from Ferguson’s Full-Scale IQ score of 77 yields an adjusted score of 73.8251, but not the product of 73.4 to which Dr. Shaffer testified.

⁶⁶ See doc. no 47 (Petitioner’s Brief), at 15 (“At the hearing, Dr. King agreed that ‘reporting an IQ score with an associated confidence interval is a critical consideration [*sic: King actually said ‘I think it’s important’*] underlying the appropriate use of intelligence tests and best practices.’” (citing doc. no. 46 (Hearing Transcript), at 180, lines 7-10).

⁶⁷ Doc. no. 46 (Hearing Transcript), at 119-20.

Dr. Shaffer administered the WAIS-IV assessment instrument to Ferguson on September 18, 2017. Dr. King re-administered the same test just five months later, on February 13, 2018, but made no adjustment for the “practice effect,” even though he acknowledged that an IQ test score can be inflated “when a test is given twice within a fairly short period of time,” ordinarily no more than six months apart, because the individual “can remember some of the items and perhaps have an increased performance on the subsequent administration of that test.”⁶⁸ Dr. King explained his failure to make any adjustment to Ferguson’s IQ score on the WAIS-IV by pointing to the fact that his Wechsler and Stanford Binet test scores were so close to one another.⁶⁹ He characterized them as in “absolute agreement.”⁷⁰

Even so, a 2012 study of the effect on IQ scores achieved by fifty-four persons to whom the WAIS-IV was re-administered at three or six months intervals after the initial, baseline administration of that test found that their Full-Scale IQ scores increased an average of 7 points. Such an increase occurred regardless of whether the re-administration occurred at three or six month intervals after the baseline test.⁷¹ In

⁶⁸ *Id.* at 138. *See also* doc. no. 47 (Petitioner’s Brief), at 15 (“Intelligence tests are made up of certain tasks, and a tester’s score is determined based on how well they perform the tasks”; and, “with practice, any task becomes easier, resulting in an upward skew known as the ‘practice effect.’”) (emphasis supplied).

⁶⁹ Doc. no. 46 (Hearing Transcript), at 139.

⁷⁰ *Id.* at 122.

⁷¹ Eduardo Estevis, Michael R. Basso, & Dennis Combs, *Effects of Practice on the Wechsler Adult Intelligence Scale Across 3- and 6-Month Intervals*, 26 Clinical Neuropsychologist 239 (Feb.

fairness, therefore, the Full-Scale IQ score achieved by Ferguson on Dr. King's re-administration of the WAIS-IV will be reduced by 7 points.

d. The majority of Ferguson's IQ scores are above 70

Ferguson concedes that "a few" of his "IQ test result ranges are above 70, even when accounting for the Flynn effect and the SEM," but he contends that "the majority of [his] tests have ranges that place his IQ scores below 70, which is strong evidence of significantly subaverage intelligence."⁷² He argues that, because "most" of his IQ scores "fall 'within the clinically established range for intellectual-functioning deficits,' the Court must 'continue the inquiry and consider other evidence of intellectual disability.'"⁷³ He contends that he has satisfied the first prong of the *Atkins* test for intellectual disability, because the "neurocognitive tests and the totality of [his] IQ test results show that his intelligence is significantly subaverage."⁷⁴

The following chart sets out the dates on which IQ assessment instruments were administered to Ferguson, the standardized tests administered on each occasion,

2012). The participants' Verbal Comprehension, Working Memory, Perceptual Reasoning, Processing Speed, and General Ability IQ sub-test indices also increased 5, 4, 5, 9, and 6 points, respectively.

⁷² Doc. no. 47 (Petitioner's Brief), at 18.

⁷³ *Id.* at 19 (citing *Moore v. Texas*, 137 S. Ct. 1039, 1050 (2017)). **Nota bene:** *Moore v. Texas* is not retroactive to cases on collateral review. See *In re Bowles*, 935 F.3d 1210, 1220 (11th Cir. 2019); *Smith v. Comm., Ala. Dep't of Corr.*, 924 F.3d 1330, 1337-40 (11th Cir. 2019).

⁷⁴ Doc. no. 47 (Petitioner's Brief), at 25.

the Full-Scale IQ score he achieved on each, followed by his scores after adjustment for the Flynn effect, and then the standard error of measurement.

Date of Test	IQ Test Given	Test Normed	Full-Scale Score	Adjusted for Flynn effect	SEM Range
1979	SB-3	1973	77	75.2	70.2 - 80.2
1985	WISC-R	1974	71	67.7	62.7 - 72.7
1988	WISC-R	1974	87	82.2	77.2 - 87.2
1997 (Dr. Rosen)	WAIS-R	1981	69	64.2	59.2 - 69.2
2017 (Dr. Shaffer)	WAIS-IV	2006	77	74.3 ⁷⁵	69.3 - 79.3 ⁷⁶
2018 (Dr. King)	WAIS-IV SB-5	2006 2003	78 ⁷⁷ 77 ⁷⁸	75.15 ⁷⁹ 72.6 ⁸⁰	70.15 - 80.15 67.6 - 77.6

⁷⁵ See the discussion in the textual paragraph accompanying notes 64 and 65, *supra*.

⁷⁶ Dr. Shaffer testified to a bidirectional SEM range of 69.4 to 78.4 (see doc. no. 46 (Hearing Transcript), at 28), but adjustment of the Full-Scale IQ score of 77 achieved by Ferguson on the WAIS-IV for the Flynn effect reduced that number to 74.3 (see the discussion in the textual paragraph accompanying notes 64 and 65, *supra*), and required a re-computation of the SEM's bidirectional range.

⁷⁷ Dr. King's score of 85 reduced by 7 as an adjustment for the "practice effect": *see* the discussion in Part II.A.4.c of this opinion, *supra*.

⁷⁸ Dr. King's score of 84 reduced by 7 for the "practice effect": *see id.*

⁷⁹ The difference between August 2008 (WAIS-IV published) and February 18, 2018 (re-administration of that test by Dr. King) was nine-and-a-half years. Multiplying 9.5 by the Flynn factor of 0.3 produces a product of 2.85 which, when subtracted from 78 (Ferguson's Full-Scale IQ score adjusted for the "practice effect") yields an IQ score adjusted for the Flynn effect of 75.15.

⁸⁰ The "Stanford Binet Intelligence Scales-Fifth Edition (SB-5)," as revised by Gale H. Roid, Ph.D., was released by Riverside Publishing Co. on May 20, 2003. *See* <https://www.businesswire.com/news/home/20030520005137/en/Riverside-Publishing-Announces-Edition-Stanford-Binet-Intelligence-Scales> (last visited Apr. 25, 2020). The difference between May 20, 2003 and February 18, 2018, when the SB-5 was administered to Ferguson by Dr. King, is 14.75 years (14 years and 9 months); and 14.75 multiplied by the Flynn factor of 0.3 produces 4.425 which, when subtracted from 77 (Ferguson's Full-Scale IQ score on the SB-5 adjusted for the "practice effect") yields an IQ score adjusted for the Flynn effect of 72.575, or 72.6 rounded.

The preceding chart graphically demonstrates that only two of the seven standardized tests administered between 1979 and 2018 resulted in IQ scores below 70, when adjusted for the Flynn effect: *i.e.*, the 1985 WISC-R and the 1997 WAIS-R. The record also reflects, however, that the persons who administered both of those tests remarked that Ferguson did not put forth his best effort. For example, the person who administered the 1985 WISC-R noted that, even though Ferguson “did not appear to be challenged by the more difficult items on the test,” he “gave up easily on both verbal and non-verbal items.”⁸¹ The 1997 WAIS-R was administered by Dr. C. Van Rosen while Ferguson was awaiting trial for capital murder in state court, and Dr. Rosen stated in his written report of evaluation that he considered it “quite apparent that [Ferguson] did not attempt to make a good effort in this test, giving up readily on many items and seemingly not trying as hard as possible. [Ferguson’s] intellectual functioning is consequently considered to be within the higher portion of the borderline range of abilities.”⁸² (As noted earlier, Dr. Rosen also testified at trial

⁸¹ Doc. no. 39-5 (Petitioner’s Exhibit 5), at 22. As recorded in note 13, *supra*, Ferguson argues that this court cannot use the comments of the examiner on the 1985 WISC-R test to conclude that his low score was due to malingering, because the examiner also stated that Ferguson was “cooperative during the entire session,” and that his scores were “considered to be valid.” Doc. no. 47 (Petitioner’s Brief), at 17. This court rejects that argument because Ferguson’s Full-Scale IQ scores before and after that test were so much higher.

⁸² Doc. no. 41-9 (Respondent’s Exhibit 5), at 7 (Rosen Report) (alterations supplied). (The Bates Number stamped in the lower, right-hand corner of this page of Respondent’s Exhibit 7 is 623.)

and reiterated that it was clear to him that Ferguson did not put forth full effort during administration of the WAIS-R, and that, if he had “really tried[,] he would have scored probably in the middle 70’s . . . perhaps a little higher.”⁸³)

Ferguson argues that it would be illogical and unfair to discredit his IQ test scores on the basis of such a rationale because, “[o]utside of Dr. Rosen’s speculative statements, there are no objective facts in the record to show that [he] *malingered* on this IQ test.”⁸⁴ He claims that it is “eminently possible” that he gave up easily because the tasks and questions were beyond his intellectual ability. He also notes that Dr. Shaffer administered a “Memory Malingering test” to determine whether he “was putting forth his best effort on the various tests as part of his examination.”⁸⁵ Based upon the results of that test, Dr. Shaffer “assumed” that the results of the WAIS-IV he administered to Ferguson were valid.⁸⁶

The State contends that a review of Dr. Rosen’s credentials from his penalty-phase testimony in the trial court shows that he was highly qualified on the basis of education and clinical experience to form an opinion about whether Ferguson was

⁸³ Doc. no. 41-9 (Respondent’s Exhibit 6), at 25 (Rosen Trial Testimony). (The Bates Number stamped in the lower, right-hand corner of this page of Respondent’s Exhibit 8 is 641.)

⁸⁴ Doc. no. 47 (Petitioner’s Brief), at 18 (alterations and emphasis supplied).

⁸⁵ *Id.* at 17.

⁸⁶ Doc. no. 46 (Hearing Transcript), at 41.

malingering, and that Ferguson has presented no evidence to the contrary.⁸⁷ The State adds that the results of the Memory Malingering test administered by Dr. Shaffer “have nothing to do with whether Ferguson malingered during Dr. Rosen’s evaluation in 1997 before his capital murder trial.”⁸⁸

This court is not persuaded by Ferguson’s argument. He has offered nothing to indicate that Dr. Rosen, an experienced clinical psychologist, was not qualified to state an opinion that Ferguson did not devote his full and best intellectual effort during administration of the WAIS-R prior to trial in state court, based upon Dr. Rosen’s personal observations of Ferguson, and the specific inconsistencies he observed during the testing processes. This court accepts Dr. Rosen’s subjective opinion that Ferguson’s IQ scores on that 1997 assessment instrument were not accurate representations of his intellectual abilities.⁸⁹

Ferguson next argues that “the majority” of his test scores, as adjusted by the standard error of measurement, yield “ranges that place his IQ scores below 70, which is strong evidence of significantly subaverage intelligence.”⁹⁰ That assertion is not

⁸⁷ Doc. no. 50 (Respondent’s Brief), at 37 (citing doc. no. 41-9 (Respondent’s Exhibit 8), at 13-16 (Rosen’s Trial Testimony)).

⁸⁸ *Id.* at 37-38.

⁸⁹ For the same reason stated in note 81, *supra*, this court could also reject Ferguson’s argument about Dr. Rosen’s observations, because Ferguson’s Full-Scale IQ scores before and after the 1997 WAIS-R test were consistently higher.

⁹⁰ Doc. no. 47 (Petitioner’s Brief), at 18.

accurate. Half of his Full-Scale IQ test scores — *i.e.*, the 1979 SB-3, 1988 WISC-R, and 2018 WAIS-IV — even when adjusted for the Flynn effect and the Standard Error of Measurement, have IQ ranges completely above 70. Three of his tests — *i.e.*, the 1985 WISC-R, 2017 WAIS-IV, and 2018 SB-5 — even when adjusted for the Flynn effect and the Standard Error of Measurement, have IQ ranges only partially below 70. The WISC-R test administered during August of 1985, when Ferguson was twelve years old and on the threshold of the 1985-86 school year, resulted in a SEM range of 62.7 to 72.7. However, the test administrator noted that, even though Ferguson “did not appear to be challenged by the more difficult items on the test,” he “gave up easily on both verbal and non-verbal items.”⁹¹ Further, the standard error of measurement for the test administered by Dr. Shaffer in 2017 results in a range from 69.3 to 79.3, and the lower end of that range barely dipped below 70.⁹²

Only the WAIS-R test administered by Dr. Rosen in 1997, prior to trial in state court, has an adjusted range totally below 70 — and the court previously discounted that score based upon Dr. Rosen’s opinion that the low scores were the result of Ferguson’s malingering, rather than an accurate measure of his intellectual ability.

⁹¹ Doc. no. 39-5 (Petitioner’s Exhibit 5), at 22.

⁹² As discussed in note 76, *supra*, Dr. Shaffer testified to a bidirectional SEM range of 69.4 to 78.4 (see doc. no. 46 (Hearing Transcript), at 28), but adjustment of the Full-Scale IQ score of 77 achieved by Ferguson on the WAIS-IV for the Flynn effect reduced that number to 74.3 (see the discussion in the textual paragraph accompanying notes 64 and 65, *supra*), and required the re-computation of the SEM’s bidirectional range that is stated in text.

Thus, the court finds that the majority of Ferguson’s IQ scores — even when adjusted for both the Flynn effect and the Standard Error of Measurement — are in a range above 70.

However, Ferguson’s scores on those three tests — when adjusted for the Flynn effect and the Standard Error of Measurement — indicate that it is *possible* his IQ falls within the range of intellectual disability.

Dr. Shaffer’s opinion that Ferguson suffers from significantly subaverage intellectual functioning is based on the results of Ferguson’s 1985, 1997, and 2017 tests, which exhibit adjusted IQ ranges that include scores below 70. However, Dr. Shaffer did not consider the possibility that the 1985 and 1997 ranges are artificially low because Ferguson malingered, and did not fully apply himself in taking those tests.⁹³ Moreover, Shaffer’s inclusion of the 2017 test, on which Ferguson had an adjusted IQ range of 69.3 to 79.3,⁹⁴ requires a presumption that Ferguson’s true IQ score falls at the very bottom of that range. But, the standard error of measurement “does not carry with it a presumption that an individual’s IQ falls to the bottom of his IQ range.” *Ledford*, 818 F.3d at 641.⁹⁵ Rather, in the case of the 2017 test, it is more

⁹³ Ferguson’s 1985 and 1997 IQ scores could place him in the range of intellectual disability. But, the fact that there is evidence that Ferguson was malingering on those tests makes that chance less likely.

⁹⁴ See note 76, *supra*.

⁹⁵ As the Eleventh Circuit noted, however, the standard error of measurement is a “bi-directional concept.” *Raulerson*, 928 F.3d at 1008 (citing *Ledford*, 818 F.3d at 641). “The standard

likely that Ferguson's true IQ score is above 70. Further, the results of all the other IQ tests Ferguson has taken place his range of scores above 70.

In light of the three scores placing him above the range of intellectual disability, the three scores placing him within the range of intellectual disability do not establish that Ferguson suffers from *significantly* subaverage intellectual functioning.

B. Proof of Significant Limitations in Adaptive Behavior

The clinical authorities all agree that an individual may achieve an IQ score greater than 70, but still “have such severe adaptive behavior problems . . . that the person’s actual functioning is comparable to that of individuals with a lower IQ score.”” *Hall*, 572 U.S. at 712 (quoting DSM-5, at 37). Thus, this court must determine whether Ferguson established, by a preponderance of the evidence, that he has significant limitations in at least two of the skills deemed by psychologists to be necessary for independent living: *i.e.*, “communication, self-care, home living, social skills, community use, self-direction, health and safety, functional academics, leisure, and work.” *Atkins*, 536 U.S. at 308 n.3 (citation omitted); *see also, e.g.*, *Perkins*, 851 So. 2d at 456(same); *Jenkins v. Commissioner, Alabama Department of Corrections*, 936 F.3d 1252, 1278 (11th Cir. 2019) (same).

error of measurement accounts for a margin of error both below *and* above the IQ test-taker’s score.” *Id.* (quoting *Ledford*, 818 F.3d at 640).

Such skills are generally referred to as “adaptive behaviors”—“functions that enable individuals to adjust to the environment appropriately and effectively,”⁹⁶ and they often are grouped into three categories for analytical purposes: *i.e.*, communication skills; daily living skills; and socialization skills. *Communication skills* refers to the “process by which one person transmits an idea to another person by means of spoken or written words, pictures, sign language, gestures, and non-verbal communications such as body language.”⁹⁷ *Daily living skills* refers to the typical, routine actions of daily living, such as “getting in and out of bed, dressing, eating, toileting.”⁹⁸ *Socialization skills* refers to the manner in which a person interacts with other people in the family and larger community. The label also includes “recreational activities and what we call coping skills, which has to do with how [the subject] handles stress and . . . manages difficult situations.”⁹⁹

1. Ferguson’s psychologist — *Dr. Robert D. Shaffer*

Dr. Shaffer utilized an assessment instrument known as the “Vineland Adaptive Behavior Scales” to evaluate Ferguson’s adaptive behaviors.¹⁰⁰ That test comes in

⁹⁶ *Dictionary of Psychology* 17.

⁹⁷ *Id.* at 191.

⁹⁸ *Id.* at 14.

⁹⁹ Doc. no. 46 (Hearing Transcript), at 35-36 (Shaffer Testimony) (alterations supplied). *See also* doc. no. 47 (Petitioner’s Brief), at 27 (defining the three domains).

¹⁰⁰ Dr. Shaffer said that he selected the Vineland Scales because he considered the assessment instrument to be “the gold standard, in terms of assessing adaptive behaviors.” Doc. no. 46 (Hearing Transcript), at 32.

three versions: “Two are administered through parents or care-givers, and the third is to be filled out by a classroom teacher.”¹⁰¹ In other words, all versions are designed to be administered by “an informant who was very familiar with the behavior of the person” to be evaluated.¹⁰² Dr. Shaffer administered the Vineland to petitioner’s mother, Mrs. Betty Ferguson, because she “was able to provide observations of [her son] at the time of his developmental period at age 18.”¹⁰³

The scaled Vineland scores assigned to Ferguson as a result of his mother’s responses to questions in each domain were: communication 67; daily living skills 67; and, socialization 68. His composite score was 63, which placed him in the first

¹⁰¹ *Dictionary of Psychology* 1052. Dr. Shaffer testified that the “Vineland Adaptive Behavior Scales” is an extension of the “Vineland Social Maturity Scale,” which the foregoing *Dictionary of Psychology* defines on the same page as:

A test used in assessing the development of individuals, including possible mental deficiency, from infancy to 30 years of age. *Persons acquainted with participants rate them* on self-help, locomotion, communication, self-direction, socialization, and occupation. Named after the Vineland Training School for the mentally retarded by Edgar A. Doll, the author of the scale.

Id. (emphasis supplied); *see also* doc. no. 46 (Hearing Transcript), at 32-33 (Shaffer Testimony).

¹⁰² *Id.* at 33. Dr. Shaffer testified on the same page that:

It’s not standard procedure with the Vineland to interview the person that’s being rated. Instead, it is intended to be a process of observing actual behaviors; that is, someone who’s watched the person, can talk about very specific concrete actions, and report those actions. Then those behaviors are subject to a very specific scoring rubric that’s identified in the Vineland manual that then results in a comparison with the U.S. population.

¹⁰³ *Id.*

percentile of the population,¹⁰⁴ meaning that he was “exceeded by 99 out of a hundred comparable individuals at age 18 in the development or in the demonstration of these independent living skills, behaviors that are necessary to perform daily routines.”¹⁰⁵

Dr. Shaffer testified that the description of Ferguson’s “prior work behavior, the behaviors described related to his marriage that were made available . . . through prior testimony, and the results of the Vineland Adaptive Behavior Scales” test, all combined to reinforce his conclusion that Ferguson suffered from significant deficits in adaptive behavior skills.¹⁰⁶

2. The State’s psychologist — *Dr. Glen D. King*

Dr. King testified that the American Association for Intellectual and Developmental Disabilities recommends the administration of either the “Vineland Adaptive Behavior Scales” or the “Adaptive Behavior Assessment System (ABAS)” test instrument for formal assessment of an individual’s adaptive behavior skills.¹⁰⁷

¹⁰⁴ *Id.* at 34-36.

¹⁰⁵ *Id.* at 37.

¹⁰⁶ *Id.* (alteration supplied). Dr. Shaffer added that he had also reviewed Ferguson’s special education records from the Morgan County and Hartselle City School Systems. *Id.* at 38.

¹⁰⁷ Doc. no. 46 (Hearing Transcript), at 146. Capital murder cases present special problems in assessing a defendant’s adaptive functioning abilities, as noted by Dr. King:

One problem is that when we use standardized instruments, like the Vineland or the Adaptive Behavior Assessment System, we’re basically asking somebody to try to fill out those devices based on their interaction with the person now. We have to try to do some kind of retro grade evaluation to the time of the offense or before then. We sometimes – [as] in this case, [the retrograde evaluation] can be 25, 26, years later.

Dr. King preferred to use the “Adaptive Behavior Assessment System – Third Edition (ABAS-3)”¹⁰⁸ and “Independent Living Scales (ILS)”¹⁰⁹ assessment instruments, because both allowed “self-report norms,” meaning that *the test subject* is allowed to “answer the questions based on his own knowledge of what he was capable of doing.”¹¹⁰ Ferguson’s scores on the ABAS-3 ranged from 7 to 13 in all of his adaptive skill areas. “His general adaptive composite, again based on an average of a hundred[,] was 98. His conceptual, social, and practical domains were 96, 111

The Vineland [Adaptive Behavior Scale] has a problem[,] in that it does not allow for self-report norms, meaning there’s no way for an individual to rate himself. You have to use an individual who has regular contact with that person on almost a daily basis for the previous six months. That’s the way the Vineland was normed.

There is no . . . provision for asking somebody to fill out the Vineland for 27 years ago. That violates standardization procedures.

Doc. no. 46 (Hearing Transcript), at 145-46 (alterations and emphasis supplied). Dr. Shaffer voiced the opposite criticism, saying: “It’s very easy for a person to have bias about their own specific abilities. And they can upgrade them or downgrade them depending on their particular bias. Most people tend to view their own capabilities as stronger than they actually are.” *Id.* at 34.

¹⁰⁸ The third edition of the Adaptive Behavior Assessment System was released in 2015 and, according to its publisher, the test includes five rating forms, each for a specific age range and evaluator: *i.e.*, Parent/Primary Caregiver Form (ages 0–5); Teacher/Daycare Provider Form (ages 2–5); Parent Form (ages 5–21); Teacher Form (ages 5–21); and *Adult Form* (ages 16–89). See (ABASTM-3) Adaptive Behavior Assessment SystemTM, Third Edition – Product Page, WPS, <https://www.wpspublish.com/abas-3-adaptive-behavior-assessment-system-third-edition> (last visited Apr. 10, 2020).

¹⁰⁹ For a description of the Independent Living Scales, see the website for “The Center for Outcome Measurement in Brain Injury,” *Introduction to the Independent Living Scale*, <https://www.tbims.org/ils/index.html> (last visited May 4, 2020).

¹¹⁰ Doc. no. 46 (Hearing Transcript), at 146.

and 96 respectively.”¹¹¹

Dr. King used the Independent Living Scales assessment instrument as an additional means of testing Ferguson’s adaptive functioning abilities, because “it has a number of questions and tasks . . . that have to do with everyday behavior,” and “it is normed and has standardization scores associated with an age range.”¹¹² He explained that the ILS assessment instrument

has items on it, for example, [like] showing him a bill and giving him a simulated check and asking him to fill out the check to pay the bill to see if he can do that. And . . . all of these [individual] items are scored on a [scale running from Zero to Two, with a] zero, meaning no points, one point meaning they get partial credit, or two points meaning they get to full criteria.

And so we get a score on these various domains for this test, as well. The subscales include things — include memory and orientation, managing money, managing home and transportation, health and safety, and social adjustment. And there are two factors on this test, as well, that have to do with problem solving and performance. Again, we combine all of these subscores and we kind of come up with a full scale standard score that indicates the ability to actually function independently.

Id. at 148 (alterations supplied). Ferguson performed well on the ILS. His full scale score, based on an average of 100, with a standard deviation of 15, was 98.¹¹³ Ferguson’s scores in all of the subparts of the ILS indicated good functioning, with

¹¹¹ Dr. King stated that a “significantly low” score “would be a score of three or below.” *Id.* at 147.

¹¹² *Id.* at 147-48.

¹¹³ *Id.* at 148.

the exception of social adjustment, which was not surprising in view of the number of years he had spent on death row on the date of testing.¹¹⁴

Dr. King's overall impression, based upon the results obtained from administration of the ABAS and ILS, on both of which Ferguson scored in the average, or non-impaired range, is that he does not meet the criteria for the adaptive functioning prong of the *Atkins* intellectual disability definition.¹¹⁵

3. The parties' arguments

Ferguson contends that Dr. Shaffer employed appropriate assessment instruments for assessing his adaptive behavior, and that Dr. Shaffer established that he suffers from significant or substantial deficits in his adaptive functioning abilities.¹¹⁶ He argues that, when an objective person takes into account the fact that he has been incarcerated for many years, with little regular contact with any persons other than prison guards, Dr. Shaffer's administration of the Vineland Adaptive Behavior Scales to Mrs. Betty Ferguson was the best means for determining his adaptive functioning abilities during his developmental period.¹¹⁷

Dr. King acknowledged the difficulty of obtaining out-of-prison observational

¹¹⁴ *Id.* at 148-49, 153.

¹¹⁵ *Id.* at 153.

¹¹⁶ See doc. no. 47 (Petitioner's Brief), at 25-26.

¹¹⁷ See *id.* at 29.

data for a death row inmate,¹¹⁸ but he still insisted there were “significant problems” with Dr. Shaffer’s administration of the Vineland Scales to Ferguson’s mother: *e.g.*,

(1) the Vineland is normed for people who have regular contact with the individual for the previous six months which did not occur in Ferguson’s case; (2) Ferguson’s mother was assessing his behavior for 27 years ago; and (3) Ferguson’s mother was not around Ferguson for the [entire] time she was evaluating him (when he was 18) because he was not living with her during this time and, in fact, was estranged from her.

Doc. no. 50 (Respondent’s Brief), at 49 (citing the testimony of Dr. King found at doc. no. 46 (Hearing Transcript), at 154). The State argues that, due to such problems, the results obtained from Dr. Shaffer’s administration of the Vineland Adaptive Behavior Scales to Ferguson’s mother is “suspect and should not be considered by this Court.” *Id.*

In response, Ferguson’s brief contends that Dr. Shaffer’s personal observations of Ferguson’s adaptive functioning abilities during the many hours that Dr. Shaffer spent with him supports the results obtained from administration of the Vineland test to his mother:

Over the course of his examination, Dr. Shaffer met with Mr. Ferguson on three separate occasions and interacted with him for nearly fourteen hours. [See doc. no. 46 (Hearing Transcript), at] 11. Mr. Ferguson is a death row inmate who has been in segregated confinement since 1998 and Dr. Shaffer is likely among a handful of individuals with the most,

¹¹⁸ See doc. no. 46 (Hearing Transcript), at 184 (“The problem . . . is that the only people who might know him well are guards, and I’m not going to ask them about him.”).

and may be the person with *the most*, significant social interaction with Mr. Ferguson in the past two decades.^[119] *See* [doc. no. 41-4 (Part 2 of Respondent's Exhibit 4), at 33 (noting, in a report dated February 16, 2019], that Mr. Ferguson had been in segregation since September 8, 1998). In his extensive observations of Mr. Ferguson, Dr. Shaffer observed behavior that was entirely consistent with the Vineland results.^[120] [See doc. no. 46 (Hearing Transcript), at] 37.

Doc. no. 47 (Petitioner's Brief), at 28-29 (*italicized* emphasis in original, citation alterations and footnotes supplied).

Ferguson also claims that the results obtained by Dr. Shaffer's administration of the Vineland to his mother were "amply supported" by evidence from his 1985 public school records which show that he was eligible for special services as an "educationally mentally handicapped" student and, therefore, could not function independently.^[121] That was "a restrictive and structured educational setting," with students staying in the same room all day with only six or seven others.^[122] Even within that close setting, one of Ferguson's teachers noted that he lacked "the

^[119] Petitioner's brief includes a footnote at the end of this sentence stating: "To contrast, Dr. King spent approximately five hours with Mr. Ferguson for his evaluation. [See doc. no. 41-1 (Respondent's Exhibit 1),] at 3."

^[120] Contrary to Ferguson's representation in the last sentence of this quotation from his Brief, Dr. Shaffer *did not describe* "observed behavior that was *entirely consistent with* the Vineland results." Instead, he testified that, when he met with Ferguson, he did not observe anything that appeared to him to be "*inconsistent with* the results of the Vineland-2 test" administered to Mrs. Betty Ferguson. Doc. no. 46 (Hearing Transcript), at 37, lines 9-11 (emphasis supplied).

^[121] *See* doc. no. 47 (Petitioner's Brief), at 28.

^[122] *Id.* (citing doc. no. 46 (Hearing Transcript), at 127).

motivation to . . . work on his own" and required "individual attention."¹²³

The State contends that Ferguson fails to acknowledge that, when he achieved a Full-Scale IQ score of 87 on the WISC-R in March of 1988, he was removed from classes for "educationally mentally handicapped" students and reassigned to the less restrictive environment of classes for "learning disabled" students.¹²⁴ The examiner who administered the assessment instrument commented that Ferguson's score placed him in the "low average range of intellectual functioning."¹²⁵ Ferguson admits the reclassification, but argues the fact that he

struggled with school after being moved to a less structured setting is further evidence of his deficits in adaptive functioning. According to Dr. Chudy, once Mr. Ferguson was reclassified as learning disabled, Mr. Ferguson "found it increasingly difficult to do the work." [Doc. no. 41-9 (Respondent's Exhibit 7), at 45 (Chudy Written Report)]. Mr. Ferguson failed most of his classes, "particularly those in which he was placed in a regular classroom setting." *Id.* In the classes with a lot of special education support, Mr. Ferguson earned "average to above average scores." *Id.* In other words, Mr. Ferguson could not independently function in a general-setting classroom like typical students. After he was removed from the EMR classroom, Mr. Ferguson "found the work to be progressively more difficult" and dropped out of

¹²³ Doc. no. 39-5 (Petitioner's Exhibit 5), at 48 (April 16, 1987 letter handwritten by Ms. Mary Beth Henry, who apparently was Ferguson's Home Economics teacher. The relevant portion read as follows: "Dale's frequent absences have put him behind on several occasions and Dale lacks the motivation to make up his work on his own. Dale works best when I give him individual attention.").

¹²⁴ Doc. no. 50 (Respondent's Brief), at 47-48. The test was required by public policy: "Dale is currently enrolled in the special education program *and is due the required three year re-evaluation for continued placement in that program.*" Doc. no. 39-5 (Petitioner's Exhibit 5), at 51 (emphasis supplied).

¹²⁵ Doc. no. 39-5 (Petitioner's Exhibit 5), at 52.

school. *Id.*

Doc. no. 47 (Petitioner's Brief), at 28.¹²⁶ The State contends that the foregoing argument is a misrepresentation of the record, and argues that Ferguson "was successful in school" after reassignment to the less restrictive environment of classes for learning disabled students.¹²⁷

Unless this court has misread the documentary evidence, however, neither Ferguson nor the State is entirely correct.

To begin, Ferguson was re-tested with the "Wechsler Intelligence Scale for Children-Revised (WISC-R)" on March 1, 1988, which would have been during the second semester of his 8th Grade, 1987-88 school year. During that semester, while still enrolled in the restrictive, small-group classroom environment for educationally mentally handicapped students, Ferguson's academic performance was acceptable, but not outstanding: that is, in January 1988 (*i.e.*, the first month of the second

¹²⁶ Dr. James F. Chudy included the following observations in his 1998 state-court report:

Mr. Ferguson continued in school until he was nearly 17 years-old and was approaching the 10th grade. At that time they had placed him in learning disabilities class and he was finding it increasingly difficult to do the work. As his school work reveals he was failing most of his classes, particularly those in which he was placed in a regular classroom setting. In those classes where he got a-lot [*sic*] of special educational help he was earning average to above average scores. In any case he found the work to be progressively more difficult, so he quit school.

Doc. no. 41-9 (Respondent's Exhibit 7), at 45 (Chudy Written Report).

¹²⁷ Doc. no. 50 (Respondent's Brief), at 48.

semester of the 1987-88 school year, and prior to the March 1, 1988 WISC-R re-test), Ferguson achieved passing grades of “C” or better in Reading, Math, Language Arts/English, Social Studies, Science, Spelling, and Handwriting. Physical Education was the only class in which Ferguson did not achieve a grade of “C” or better.¹²⁸

However, following his transfer to the less structured classes for “learning disabled” students at the beginning of the following, 1988-89 school year, Ferguson faltered. His grades for the first semester of that 9th Grade school year were three D’s (Special Reading, Physical Education, and Agricultural Business I) and three F’s (Special Math, Special Social Studies, and Special Science).¹²⁹ His grades for the second semester of the 9th Grade, 1988-89 school year improved, but only somewhat. He earned one “B+” (Reading 9-12), one “D” (Science 9/L), two “D-minuses” (Math 9/L and Physical Education), and two “F’s” (Social Studies and Agricultural Business I).¹³⁰

Then, inexplicably, Ferguson’s academic performance in both semesters of the following 1989-90, 10th Grade school year improved *significantly*. During the first semester, he earned two “A’s” (Math GED and History T/L), two “B’s” (English T/L

¹²⁸ *Id.* (citing doc. no. 39-5 (Petitioner’s Exhibit 5), at 84).

¹²⁹ See doc. no. 39-5 (Petitioner’s Exhibit 5), at 86 (“Grade 9, 89 Year, 1 Sem.”); *see also id.* at 87 (a second, but less legible, copy of the same page from Ferguson’s school records).

¹³⁰ *Id.* (“SM: 2 YR: 88-89”).

and Science T/L), one “C- minus” (Physical Ed/B), and one “F” (Health Occ I).¹³¹ During the second semester, he earned *three* “A’s” (English T/L, History T/L, and Science T/L), one “C- minus” (Physical Ed), one “D” (Math GED), and one “F” (Health Occ I).¹³² It only was during the first semester of the 1990-91, 11th Grade school year that Ferguson stumbled badly, yet again, and received three “F’s” in English TL, Math TL, and Apartment Living, and three “Incompletes” in Math TE & TL, History TL, and Health/Science.¹³³ He thereafter withdrew from school and his “Incompletes” were changed to “F’s.”

The bottom line is: the official Hartselle City Schools’ record of Ferguson’s academic performance following his removal from classes for “educationally mentally handicapped” students and reassignment to the less restrictive environment of classes for “learning disabled” students is a mixed bag into which both parties can reach to find support for their respective arguments. Even so, there is more support for the State’s position than for Ferguson’s.

The State also argues that the school records relied upon by Ferguson are not consistent with the results obtained from Dr. King’s administration of the “Wide Range Achievement Test – Fourth Edition (WRAT-4).” When Ferguson was tested

¹³¹ *Id.* (“SM: 1 YR: 89-90”).

¹³² *Id.* (“SM: 2 YR: 89-90”).

¹³³ *Id.* (“SM: 1 YR: 90-91”).

by Dr. King in 2018, he was reading at a 9.5 grade level, spelling at a 7.5 grade level, and performing math calculations at a 6.1 grade level.¹³⁴ Dr. King believes that the ability to read and spell at those levels is not consistent with a contention that Ferguson is intellectually disabled.¹³⁵ This court agrees.

Shifting focus, Ferguson argues that neither of the assessment instruments administered by Dr. King — *i.e.*, the “Adaptive Behavior Assessment System” and “Independent Living Scales” — returned accurate or reliable data because, “[c]ontrary to best practices outlined in the authoritative texts,” both of those tests relied on Ferguson’s “self-reporting about whether he believed he was able to do real-life tasks,” and Dr. King did not consider how the “results may be inconsistent with the record.”¹³⁶ Ferguson criticizes the Adaptive Behavior Assessment System, “in which Dr. King asked Mr. Ferguson to rate his own abilities on whether he could perform a number of tasks,” because most of the questions pertained to tasks Ferguson would never have an opportunity to perform in his prison environment, such as: making appointments by telephone, mobile device, or internet; using paper or digital maps to find his way to desired locations; calling a repairman when needed; showing responsibility for personal finances; or walking or riding a bike alone to

¹³⁴ Doc. no. 46 (Hearing Transcript), at 142-43.

¹³⁵ *Id.* at 144-45.

¹³⁶ Doc. no. 47 (Petitioner’s Brief), at 29-30.

locations within a mile of home or work.¹³⁷ The State responds by arguing that all of those questions addressed activities that Ferguson could have performed prior to his incarceration.¹³⁸ (The State’s response requires more credulity of this judicial officer than he is inclined to grant, however, especially when one considers that mobile telephone devices, computers, “the internet,” and “digital maps” (which are dependent upon Global Positioning Satellites) were technological innovations that either had not been perfected, or were not in widespread use, on the date Ferguson was arrested and incarcerated for the underlying capital offenses.)

Ferguson also faults Dr. King for allowing him to read the instructions for the Adaptive Behavior Assessment System test, without making any “further attempt to check [his] understanding of the test instructions beyond generally asking if [he] had any questions.”¹³⁹ He argues that the results of that test indicate its self-reporting bias.

On the ABAS-3, Mr. Ferguson gave himself the highest score on almost every single question, indicating that he “always (or almost always) when needed” could complete the tasks. *See generally* Pet’r Ex. 10 [sic]. Among the things that Mr. Ferguson indicated he could *always* do when needed were: use electrical sockets safely; check bank or other financial statements to make sure they are correct; pay bills on time; plan ahead for fun activities on free days or afternoons; and invite others to join him in playing games and other fun activities. *See generally id.*

¹³⁷ *Id.* at 30.

¹³⁸ *See* doc. no. 50 (Respondent’s Brief), at 46.

¹³⁹ Doc. no. 47 (Petitioner’s Brief), at 30 (alterations supplied).

The absurdity of these answers, considering Mr. Ferguson's highly restrictive incarceration, shows that Dr. King's administration of the ABAS-3 was plagued with self-reporting bias. *See* [doc. no. 46 (Hearing Transcript), at] 34 ("It's very easy for a person to have bias about their own specific abilities."); *id.* at 42 (Dr. Shaffer's testimony that the ABAS-3, when "applied directly to the person themselves" is "very subject to their personal bias"); *[United States v.] Lewis*, [No. 1:08 CR 404,] 2010 WL 5418901, at *23 [(N.D. Ohio Dec. 23, 2010)] ("[T]he use of Defendant as an informant to his own adaptive behavioral capabilities, known as 'self-reporting,' without any corroborating sources, is disfavored by the AAIDD."). The potential for self-reporting bias was only magnified when the test asked Mr. Ferguson to self-report on whether he could do an activity that has had no connection to his life in prison for over twenty years.

Doc. no. 47 (Petitioner's Brief), at 30-31 (alterations supplied). Ferguson contends that the mere fact he "filled out the ABAS-3 in such a manner without asking for clarification is indicative of his lack of adaptive functioning skills."¹⁴⁰ He argues that his limitations in his adaptive functioning abilities are self-evident: otherwise, he "would certainly have asked some clarifying questions to ensure that [he was] providing accurate results," and his failure to "raise any concerns" about the test questions underscores the inaccuracy of its results.¹⁴¹

The State responds by noting that there is no *evidence* in the record indicating that Ferguson did not understand the instructions for taking the ABAS-3, and contends that he "would have understood when he was rating himself that the

¹⁴⁰ *Id.* at 31.

¹⁴¹ *Id.*

questions involved activities that he had done before he was incarcerated — such as riding a bike or calling a repairman.”¹⁴²

The State also disputes Ferguson’s contention that Dr. King did not conduct a meaningful assessment of his adaptive functioning abilities by pointing to Dr. King’s uncontradicted testimony that the Adaptive Behavior Assessment System is one of two instruments recognized by the American Association on Intellectual and Developmental Disabilities for assessing a person’s adaptive functioning abilities.¹⁴³ Dr. King explained his reasons for using the ABAS as follows:

Well, again, the American Association for Intellectual and Developmental Disabilities recommends for formal assessment of adaptive behavior that either the Vineland or the ABAS [be] used. Those are acceptable instruments to determine adaptive functioning when you’re trying to come up with some standardized scoring.

I prefer the Adaptive Behavior Assessment System for the reasons I’ve already mentioned because it has self-report norms so the individual can take it now and answer the questions based on his own knowledge of what he was capable of doing.

Doc. no. 46 (Hearing Transcript), at 146 (alteration supplied).

Additionally, the State disputes Ferguson’s contention that the ABAS and Independent Living Scales are contrary to the so-called “best practices” outlined in “authoritative texts,” because each test relies upon self-reporting. As the State notes,

¹⁴² Doc. no. 50 (Respondent’s Brief), at 46-47.

¹⁴³ *Id.* at 46 (citing doc. no. 46 (Hearing Transcript), at 146).

Ferguson did not “identify these authoritative texts or cite this Court to where these authoritative texts state that self-reporting is improper.”¹⁴⁴

Ferguson also contends that, even though Dr. King was “aware of the possibility that the ABAS-3 results were skewed and invalid,” he made no attempt to “follow up,” to determine whether Ferguson “could actually do the activities he said he could.”¹⁴⁵

Ferguson also questions the validity and usefulness of the Independent Living Scales test, which was administered by Dr. King to obtain “a measure of [his] practical abilities in Managing Money, Managing Home and Transportation, Health and Safety, Social Adjustment, and Problem Solving.”¹⁴⁶ Ferguson contends that: the assessment instrument has not been updated since 1998; it has never been approved by the AAIDD to diagnose intellectual disabilities; and the tasks and questions on the test had very little applicability to the everyday-life of a death row inmate.¹⁴⁷

The State replies to that argument by saying that, even though the Independent

¹⁴⁴ *Id.* at 46.

¹⁴⁵ Doc. no. 47 (Petitioner’s Brief), at 31-32. The petitioner’s attorneys conclude that the results of the ABAS-3 were not valid because Ferguson was required to “do an enormous amount of guesswork on whether he could perform certain tasks.” *Id.* at 32.

¹⁴⁶ Doc. no. 41-1 (Respondent’s Exhibit 1), at 9.

¹⁴⁷ Doc. no. 47 (Petitioner’s Brief), at 32. Even so, Ferguson points out that the test tasks he successfully performed (such as filling out a check and counting change) “are not even dispositive for a finding of intellectual disability,” and Dr. King conceded that intellectually disabled individuals do not need to be “completely unable to do ordinary living tasks.” *Id.* (quoting doc. no. 46 (Hearing Transcript), at 199)).

Living Scales is not an assessment instrument approved by the AAIDD, “there was no testimony that Dr. King’s use of the ILS was not appropriate”; instead, Dr. Shaffer testified that he did not administer the ILS to Ferguson “because it is his preference to use the Vineland.”¹⁴⁸ The State adds that, even though Ferguson is critical of the ILS because individuals who may be intellectually disabled can count change and complete bank checks, “Ferguson was able to do much more than fill out a check on his administration of the ILS.”¹⁴⁹

Ferguson also contends that Dr. King’s opinion that his adaptive functioning was “well within the average range” was contradicted by other evidence. He argues that he

struggled significantly in school with learning in a non-structured setting, and ultimately dropped out after he was moved to a regular classroom. *See* [doc. no. 41-9 (Respondent’s Exhibit 7)], at 45 (Dr.

¹⁴⁸ Doc. no. 50 (Respondent’s Brief), at 47; *see also* doc. no. 46 (Hearing Transcript), at 44.

¹⁴⁹ Doc. no. 50 (Respondent’s Brief), at 47. Dr. King testified that Ferguson was shown a card that has a social security check amount on it. He [was] asked to record the amount on a piece of paper. He did that.

And he was shown a bill from a gas company and was asked to fill out a check to pay that bill. He filled out the check accurately with the date, the amount, signing it. And he did that for two separate bills.

He then was asked to take the amounts of those two bills and subtract them from the social security check that he was supposedly given. And he came up with exactly the right amount.

Doc. no. 46 (Hearing Transcript), at 151 (alteration supplied).

Chudy's ~~testimony pertaining to~~ [sic: actually, it was Chudy's 1998 state-court written report summarizing] Mr. Ferguson's school performance after being reclassified). Mr. Ferguson's ex-wife testified at trial that during their marriage, "Dale did not understand things well and . . . I would make all of our decisions." *Id.* at 80. Dr. Chudy's expert report also noted Mr. Ferguson's significant social limitations:

Being intellectually slow with poor social skills he seems peculiar among others. This sets him up to be taken advantage of. . . . When he does develop friendships, his naïve and limited thinking allows him to be easily influenced. Family members confirm that in relationships where he has come to trust the other person, he is easily led and becomes a follower. . . .

Doc. no. 47 (Petitioner's Brief), at 32-33 (quoting doc. no. 41-9 (Respondent's Exhibit 7), at 48-49 (Chudy Written Report)) (footnote omitted).¹⁵⁰ Ferguson contends that the "clear inconsistencies" between the results of the assessment instruments administered by Dr. King and the record should have led Dr. King to make additional inquiries into his adaptive functioning abilities.¹⁵¹

In summary, Ferguson contends that "the adaptive functioning evidence as a whole" weighs in favor of a conclusion that he has significant and substantial limitations in his adaptive functioning abilities.

¹⁵⁰ The omitted footnote was at the end of the extract from Dr. Chudy's 1998 state-court report, and argues that his quoted description of Ferguson was "entirely in line with the Supreme Court's characterization of individuals with intellectual disability in *Atkins*. See 536 U.S. at 317 ('[T]here is abundant evidence that [individuals with intellectual disability] often act on impulse rather than pursuant to a premeditated plan, and that in group settings they are followers rather than leaders.')" Doc. no. 47 (Petitioner's Brief), at 33 n.12.

¹⁵¹ *Id.* at 33.

Dr. Shaffer used observer evidence on the Vineland test, which was consistent with his own observations and the record as a whole, to determine that Mr. Ferguson was significantly limited in adaptive functioning. On the other hand, Dr. King relied on two tests that he knew to be error-prone and invalid to form his opinion that Mr. Ferguson had average adaptive functioning skills and did not confirm his findings with any other corroborative evidence. Even a cursory glance through the record shows that Mr. Ferguson was not an “average” person when it came to adaptive functioning.

Doc. no. 47 (Petitioner’s Brief), at 34.

4. Analysis

The adaptive behavior component of an intellectual disability determination requires proof of, among other things, “substantial *present* limitation in at least two of the following areas: ‘communication, self-care, home living, social skills, community use, self-direction, health and safety, functional academics, leisure, and work.’” *Jenkins*, 936 F.3d at 1278 (quoting *Atkins*, 536 U.S. at 308 n.3) (emphasis supplied). This court is not persuaded by a preponderance of the evidence that Ferguson has made such a showing.

Although the State contends that “Ferguson has not shown that he suffers now, or has ever suffered, from substantial or significant deficits in adaptive functioning,”¹⁵² the parties’ arguments focus only upon Ferguson’s adaptive functioning abilities or deficits prior to his conviction in 1998. Ferguson’s expert, Dr.

¹⁵² Doc. no. 50 (Respondent’s Brief), at 45.

Shaffer, administered the Vineland Adaptive Behavior Scales assessment instrument to Ferguson's mother, saying that Mrs. Betty Ferguson would be able to "provide observations of [her son's adaptive functioning abilities] at the time of his developmental period at age 18."¹⁵³ Although Dr. Shaffer testified that his observations of Ferguson on the three dates that he evaluated him in preparation for testifying in this court were "not inconsistent" with the results obtained by his administration of the Vineland Scales assessment instrument to Ms. Ferguson, he offered little evidence to support a finding that Ferguson *presently* suffers from substantial deficits in any area of adaptive functioning. Further, the school records relied upon by Ferguson address only his adaptive functioning abilities decades ago, and not presently.

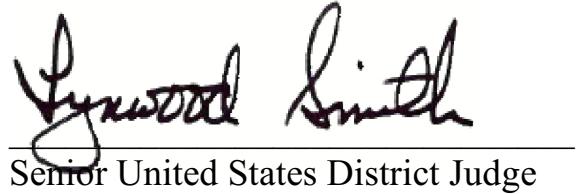
III. CONCLUSION

After consideration of the testimony and documentary exhibits presented at the August 27, 2019 hearing and the parties' post-hearing briefs, this court concludes that petitioner, Thomas Dale Ferguson, has failed to establish by a preponderance of the evidence that he suffers from significantly subaverage intellectual functioning, or that he has significant or substantial deficits in adaptive behavior. Therefore, Ferguson's *Atkins* claim and his petition for writ of habeas corpus are due to be denied. A

¹⁵³ Doc. no. 46 (Hearing Transcript), at 33.

separate final judgment will be entered contemporaneously herewith.

DONE this 21st day of May, 2020.



Lynnwood Smith
Senior United States District Judge